

METHOD AND DEVICE FOR PREPARING SUMMARY SENTENCE AND STORAGE MEDIUM STORING SUMMARY SENTENCE PREPARATION PROGRAM

Publication number: JP2001052032

Publication date: 2001-02-23

Inventor: SEKI YOSHIHIRO; IWASE SHIGETO

Applicant: NIPPON TELEGRAPH & TELEPHONE

Classification:

- international: G06F17/30; G06F17/30; (IPC1-7): G06F17/30

- European:

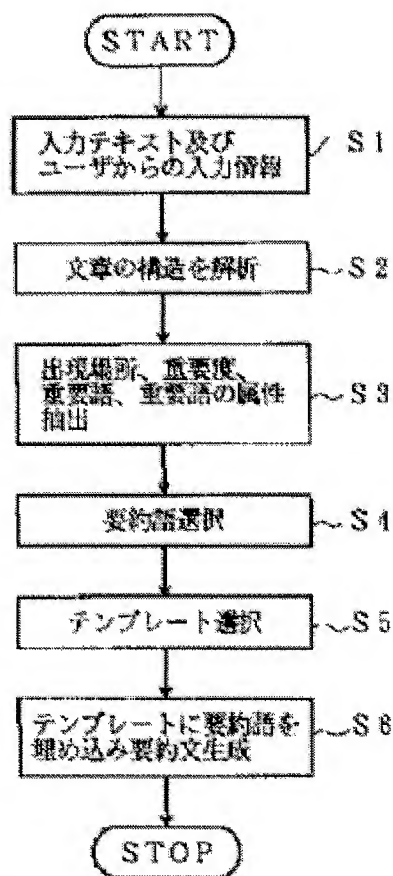
Application number: JP19990299860 19991021

Priority number(s): JP19990299860 19991021; JP19990150571 19990528

Report a data error here

Abstract of JP2001052032

PROBLEM TO BE SOLVED: To increase the choices of information for a user by selecting the template of a sentence to be generated corresponding to the attribute of a selected summary word and generating a summary sentence by embedding the summary word in that template. **SOLUTION:** In a step 1, an input text including a title, a text and charts or input information from the user is inputted and in a step 2, the structure of the inputted input text or input information is analyzed. In a step 3, the appearance places of words in the input text or input information, an importance degree, important words and the attribute of that importance degree or the like are extracted while using an important word definition table defining the importance degree of a word on the basis of the appearance position and the modification relation of the word and in a step 4, the summary word to be used for the summary sentence is selected by weighing the extracted important word corresponding to the place of appearance and the degree of importance. In a step 5, the template of the sentence to be generated is selected corresponding to the attribute of the selected summary word and in a step 6, the summary sentence is generated by embedding the summary word in the selected template.



Data supplied from the esp@cenet database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2001-52032
(P2001-52032A)

(43) 公開日 平成13年2月23日 (2001.2.23)

(51) Int.Cl.⁷
G 0 6 F 17/30

識別記号

F I
G 0 6 F 15/401
15/40

テ-マコ-ト*(参考)
3 2 0 A 5 B 0 7 5
3 7 0 A

審査請求 未請求 請求項の数30 O L (全 25 頁)

(21) 出願番号 特願平11-299860

(22) 出願日 平成11年10月21日 (1999.10.21)

(31) 優先権主張番号 特願平11-150571

(32) 優先日 平成11年5月28日 (1999.5.28)

(33) 優先権主張国 日本 (J P)

(71) 出願人 000004226

日本電信電話株式会社
東京都千代田区大手町二丁目3番1号

(72) 発明者 関 良博

東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内

(72) 発明者 岩瀬 成人

東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内

(74) 代理人 100070150

弁理士 伊東 忠彦

Fターム(参考) 5B075 ND03 NS01 UU24

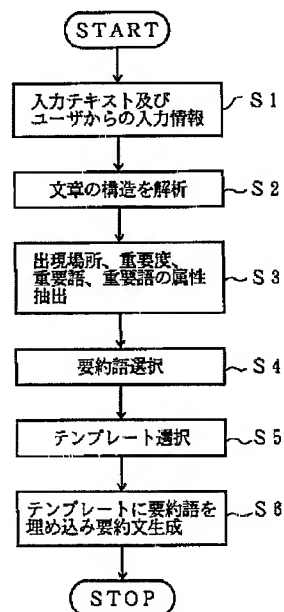
(54) 【発明の名称】 要約文作成方法及び装置及び要約文作成プログラムを格納した記憶媒体

(57) 【要約】

【課題】 有用な言葉のみで要約文を作成することが可能な要約文作成方法及び装置及び要約文作成プログラムを格納した記憶媒体を提供する。

【解決手段】 本発明は、入力された入力テキストのタイトル、本文、表を含む文章やユーザからの入力情報の構造を解析し、出現位置と単語の係り受け関係で単語の重要度を定義した重要単語定義テーブルを用いて単語の出現場所、重要度、重要語、該重要語の属性を抽出し、抽出した重要語に対し、出現場所と重要度で重み付けをして要約で用いる要約語を選択し、選択された要約語の属性に応じて生成する文のテンプレートを選択し、選択されたテンプレートに要約語を埋め込んで要約文を生成する。

本発明の原理を説明するための図



【特許請求の範囲】

【請求項1】 テキスト処理における要約文を生成する要約文生成方法において、

タイトル、本文、表を含む入力テキストや、ユーザからの入力情報を入力し、

入力された前記入力テキストや前記入力情報の構造を解析し、

出現位置と単語の係り受け関係で単語の重要度を定義した重要単語定義テーブルを用いて、前記入力テキストや前記入力情報の単語の出現場所、重要度、重要語、該重要語の属性を抽出し、

抽出した前記重要語に対し、前記出現場所と前記重要度で重み付けをして要約文に用いる要約語を選択し、

選択された前記要約語の属性に応じて生成する文のテンプレートを選択し、

選択されたテンプレートに前記要約語を埋め込んで要約文を生成することを特徴とする要約文生成方法。

【請求項2】 前記要約語を選択する際に、

常に案内する情報や、前記ユーザからの入力情報と重複する単語を含まないような重要語を選択する請求項1記載の要約文生成方法。

【請求項3】 前記重要語を選択する際に、

前記常に案内する情報や、前記ユーザからの入力情報と重複する単語を含む単語の重要度を下げる、または、フィルタリングによりカットする請求項2記載の要約文生成方法。

【請求項4】 前記要約語を選択する際に、

前記入力された前記入力テキスト及び前記入力情報について、他の文章と違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する請求項1記載の要約文生成方法。

【請求項5】 前記差分を強調するような要約文を生成するための要約語を選択する際に、

入力された複数の入力テキストや前記入力情報中に同様の表現を行う単語が出現している場合に、該単語の重要度を下げる、または、格情報、品詞情報の出現を抑止する請求項4記載の要約文生成方法。

【請求項6】 前記要約語を選択する際に、

常に案内する情報や、前記ユーザからの入力情報と重複する単語を含まないような重要語を選択すると共に、他のテキストと違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する請求項1記載の要約文生成方法。

【請求項7】 前記要約語を選択する際に、

複数のテキスト間において他のテキストと同じ属性の要約語を選択し、該要約語に属性を付与し、該属性の共通

部分を探索し、同じ表現を纏めて要約文を生成することが可能な要約語を選択し、

選択された前記要約語を用いて要約文を生成する請求項1記載の要約文生成方法。

【請求項8】 前記要約語を選択する際に、

ユーザから入力された前記入力情報から該ユーザの意図をユーザ意図推論知識を用いて推論し、該推論により抽出された単語に関連する該入力情報内の単語の重要度を上げる請求項1記載の要約文生成方法。

【請求項9】 予め画面上に表示可能な文字数や行数を含む表示領域情報を保持しておき、

前記要約文を生成した後に、該要約文と前記表示領域情報を比較し、

前記要約文が前記表示領域情報に示される領域に対応するように、該要約文を編集する請求項1記載の要約文生成方法。

【請求項10】 前記要約文を編集する際に、

前記要約文が前記表示領域情報に示される所定の範囲を越えないように文字数を調整する、または、格情報に対応した表示情報を選択する請求項9記載の要約文生成方法。

【請求項11】 テキスト処理における要約文を生成する要約文生成装置であって、

タイトル、本文、表を含む入力テキストや、ユーザからの入力情報を入力する文章入力手段と、

前記文章入力手段で入力された前記入力テキストや前記入力情報の構造を解析する文章構造解析手段と、

出現位置と単語の係り受け関係で単語の重要度を定義した重要単語定義テーブルと、

前記重要単語定義テーブルを用いて、前記入力テキストや前記ユーザからの入力情報の単語の出現場所、重要度、重要語、該重要語の属性を取り出す重要語抽出手段と、

前記重要語抽出手段で抽出した前記重要語に対し、前記出現場所と前記重要度で重み付けをして要約文に用いる要約語を選択する要約語選択手段と、

選択された前記要約語の属性に応じて生成する文のテンプレートを選択するテンプレート選択手段と、

選択されたテンプレートに前記要約語を埋め込んで要約文を生成する要約文生成手段とを有することを特徴とする要約文生成装置。

【請求項12】 前記要約語選択手段は、

常に案内する情報や、前記ユーザからの入力情報と重複する単語を含まないような重要語を選択する情報抑制手段を有する請求項11記載の要約文生成装置。

【請求項13】 前記情報抑制手段は、

前記常に案内する情報や、前記ユーザの入力情報と重複する単語を含む単語の重要度を下げる、または、フィルタリングによりカットする手段を含む請求項12記載の要約文生成装置。

【請求項14】 前記要約語選択手段は、前記入力された前記入力テキスト及び前記入力情報について、他の文章と違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較手段を有する請求項1記載の要約文生成装置。

【請求項15】 前記要約語比較手段は、入力された複数の入力テキストや前記入力情報中に同様の表現を行う単語が出現している場合に、該単語の重要度を下げる、または、格情報、品詞情報の出現を抑止する手段を含む請求項14記載の要約文生成装置。

【請求項16】 前記要約語選択手段は、常に案内する情報や、前記ユーザからの入力情報と重複する単語を含まないような重要語を選択する情報抑制手段と、

他のテキストと違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較手段とを有する請求項1記載の要約文生成装置。

【請求項17】 前記要約語選択手段は、入力された複数の入力テキスト間において他のテキストと同じ属性の要約語を選択し、該要約語に属性を付与し、該属性の共通部分を探索し、同じ表現を纏めて要約文を生成することが可能な要約語を選択する手段を含み、

前記要約文生成手段は、選択された前記要約語を用いて要約文を生成する手段を含む請求項1記載の要約文生成装置。

【請求項18】 前記要約語選択手段は、ユーザから入力された前記入力情報から該ユーザの意図をユーザ意図推論知識を用いて推論し、該推論により抽出された単語に関連する該入力情報内の単語の重要度を上げる手段を含む請求項1記載の要約文生成装置。

【請求項19】 画面上に表示可能な文字数や行数を含む表示領域情報を更に有し、

前記要約文生成手段は、前記要約文を生成した後に、該要約文と前記表示領域情報を比較し、該要約文が前記表示領域情報に示される領域に対応するように、該要約文を編集する編集手段を含む請求項1記載の要約文生成装置。

【請求項20】 前記編集手段は、前記要約文が前記表示領域情報に示される所定の範囲を越えないように文字数を調整する、または、格情報に対応した表示情報を選択する手段を含む請求項19記載の要約文生成装置。

【請求項21】 テキスト処理における要約文を生成する要約文生成プログラムを格納した記憶媒体であって、タイトル、本文、表を含む入力テキストや、ユーザから

の入力情報を入力させる文章入力プロセスと、前記文章入力プロセスにより入力された前記入力テキストや前記入力情報の構造を解析する文章構造解析プロセスと、

出現位置と単語の係り受け関係で単語の重要度を定義した重要単語定義テーブルを用いて単語の出現場所、重要度、重要語、該重要語の属性を取り出す重要語抽出プロセスと、

前記重要語抽出プロセスで抽出した前記重要語に対し、前記出現場所と前記重要度で重み付けをして要約文に用いる要約語を選択する要約語選択プロセスと、選択された前記要約語の属性に応じて生成する文のテンプレートを選択するテンプレート選択プロセスと、選択されたテンプレートに前記要約語を埋め込んで要約文を生成する要約文生成プロセスとを有することを特徴とする要約文生成プログラムを格納した記憶媒体。

【請求項22】 前記要約語選択プロセスは、常に案内する情報や、前記ユーザからの入力情報と重複する単語を含まないような重要語を選択する情報抑制プロセスを有する請求項21記載の要約文生成プログラムを格納した記憶媒体。

【請求項23】 前記情報抑制プロセスは、前記常に案内する情報や、前記ユーザからの入力情報と重複する単語を含む単語の重要度を下げる、または、フィルタリングによりカットするプロセスを含む請求項22記載の要約文生成プログラムを格納した記憶媒体。

【請求項24】 前記要約語選択プロセスは、前記入力された前記入力テキスト及び前記入力情報について他の文章と違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較プロセスを有する請求項21記載の要約文生成プログラムを格納した記憶媒体。

【請求項25】 前記要約語比較プロセスは、入力された複数の入力テキストや前記入力情報中に同様の表現を行う単語が出現している場合に、該単語の重要度を下げる、または、格情報、品詞情報の出現を抑止するプロセスを含む請求項24記載の要約文生成プログラムを格納した記憶媒体。

【請求項26】 前記要約語選択プロセスは、常に案内する情報や、前記ユーザからの入力情報と重複する単語を含まないような重要語を選択する情報抑制プロセスと、

他の文書と違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較プロセスとを有する請求項21記載の要約文生成プログラムを格納した記憶媒体。

【請求項27】 前記要約語選択プロセスは、入力された複数の入力テキスト間において他のテキストと同じ属性の要約語を選択し、該要約語に属性を付与し、該属性の共通部分を探索し、同じ表現を纏めて要約文を生成することが可能な要約語を選択するプロセスを含み、前記要約文生成プロセスは、選択された前記要約語を用いて要約文を生成するプロセスを含む請求項21記載の要約文生成プログラムを格納した記憶媒体。

【請求項28】 前記要約語選択プロセスは、ユーザから入力された前記入力情報から該ユーザの意図をユーザ意図推論知識を用いて推論し、該推論により抽出された単語に関連する該入力情報内の単語の重要度を上げるプロセスを含む請求項21記載の要約文生成プログラムを格納した記憶媒体。

【請求項29】 前記要約語生成プロセスは、前記要約文を生成した後に、該要約文と画面上に表示可能な文字数や行数を含む表示領域情報とを比較し、該要約文が該表示領域情報に示される領域に対応するように、該要約文を編集する編集プロセスを含む請求項21記載の要約文生成プログラムを格納した記憶媒体。

【請求項30】 前記編集プロセスは、前記要約文が前記表示領域情報に示される所定の範囲を越えないように文字数を調整する、または、格情報に対応した表示情報を選択するプロセスを含む請求項29記載の要約文生成プログラムを格納した記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、要約文作成方法及び装置及び要約文作成プログラムを格納した記憶媒体に係り、特に、テキスト処理において、複数の電子メール、ホームページ、お店の案内文等の要約を一覧で表示するように、内容を1行程度で要約する要約文作成方法及び装置及び要約文作成プログラムを格納した記憶媒体に関する。

【0002】 詳しくは、入力テキストから文章構造を解析し、抽出された重要語の中から要約語を選択ルールのみならず、シソーラス知識に基づいて文テンプレートを選択し、それに沿った要約文を生成するための要約文作成方法及び装置及び要約文作成プログラムを格納した記憶媒体に関する。

【0003】

【従来の技術】 従来、文書を一覧表示する場合は、タイトルを表示したり、文書の先頭の文を表示している。また、要約文を作成する際、重要な文を選んで集めた抜粋文章を作る方法がある。この方法は、まず、文章の中に含まれる名詞を全て調べ、順位トップの名詞は文章の最も重要なキーワードと考えられるので、これを多く含む文ほど重要な文と考えて文毎に重要度ランキングを作

る。そして、指定された要約率に応じて低ランクの文を切捨てる。このように、従来は、要約技術に関しては重要な文を抽出する抄録のレベルが、*tf × idf* 法等による統計的手法により重要語を選択している。

【0004】

【発明が解決しようとする課題】 しかしながら、ホームページにはタイトルのない文書が多く、文書の先頭だけでは内容を表すことができない。特に、長い文を1行程度で表示した場合、途中で途切れてしまい、内容が分からない場合もある。また、キーワードがどこに出現しても同じ重みなので、タイトルに出現する単語と、末尾の補足説明にある単語が同じ重要度になってしまう。タイトルはその文章をまとめる表現であると考えられるため、そこから抽出された単語の重要度を上げる必要があるし、補足説明から抽出された単語の重要度は下げる必要がある。

【0005】 検索結果の一覧を表示する場合、各文書の内容の違いが分かるように表示する項目を統一することが有効であるが、上記従来の統計的な手法では、単語の意味をもとにキーワードを取り出していないので、「対象者」や「取扱商品」、「値段」という観点でのキーワードを取り出すことができない。また、何らかの情報と要約を同時に表示した場合、要約以外で表示する情報と重複する内容は要約に入れる必要がない。しかし、上記の従来の技術における文書一覧表示の方法では、では、そのような判断は不可能である。例えば、「高齢者を対象にした福祉サービス」で検索した場合、対象者は「高齢者」であることは明らかなので、「高齢者にヘルパーを派遣します」という要約ではなく、「ヘルパーを派遣します」で十分である。また、タイトルと要約を同時に表示した場合、タイトルと重複する内容は要約に入れる必要がない。しかし、従来の技術では、そのような判断は不可能である。

【0006】 例えば、「ヘルパー派遣」というタイトルに対しては、「ヘルパーを派遣します」という要約ではなく、それ以外の付加情報、ヘルパーの仕事内容、料金はかかるのか等の情報を表示することがユーザにとって有用である。また、「老人ホーム」で検索し、複数のサービス、例えば、「特別養護老人ホーム」「養護老人ホーム」が検索された場合、それぞれの案内サービスに対して個別に要約を作成した場合、「老人を預かり世話します」といった同一内容の情報となり、有用な要約とは言えない。検索された全てのサービスの内容を比較し、それぞれのサービスの差分を強調し、「寝たきり老人や痴呆の老人を世話します」「一人暮らし老人や低所得者層を世話します」と差分を強調し表示することによって、ユーザのサービスの選択性を向上させる必要がある。

【0007】 さらに、上記の従来の技術では、装置上の表示可能領域に基づいて動的に要約情報を変化させるこ

とができないために、例えば、案内領域が40文字文の場合、定型ロジックで要約文を編集した場合、要約文の一部しか案内することができなかつたり、あまりにも簡約な要約となり、情報が少な過ぎるといった問題が発生する。

【0008】本発明は、上記の点に鑑みなされたもので、有用な言葉のみで要約文を作成することが可能な要約文作成方法及び装置及び要約文作成プログラムを格納した記憶媒体を提供することを目的とする。

【0009】

【課題を解決するための手段】図1は、本発明の原理を説明するための図である。本発明（請求項1）は、テキスト処理における要約文を生成する要約文生成方法において、タイトル、本文、表を含む入力テキストや、ユーザからの入力情報を入力し（ステップ1）、入力された入力テキストや入力情報の構造を解析し（ステップ2）、出現位置と単語の係り受け関係で単語の重要度を定義した重要単語定義テーブルを用いて、入力テキストや入力情報の単語の出現場所、重要度、重要語、該重要語の属性を抽出し（ステップ3）、抽出した重要語に対し、出現場所と重要度で重み付けをして要約文に用いる要約語を選択し（ステップ4）、選択された要約語の属性に応じて生成する文のテンプレートを選択し（ステップ5）、選択されたテンプレートに要約語を埋め込んで要約文を生成する（ステップ6）。

【0010】本発明（請求項2）は、要約語を選択する際に、常に案内する情報や、ユーザからの入力情報と重複する単語を含まないような重要語を選択する。本発明（請求項3）は、重要語を選択する際に、常に案内する情報や、ユーザからの入力情報と重複する単語を含む単語の重要度を下げる、または、フィルタリングによりカットする。

【0011】本発明（請求項4）は、要約語を選択する際に、入力された入力テキスト及び入力情報について、他のテキストと違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する。

【0012】本発明（請求項5）は、差分を強調するような要約文を生成するための要約語を選択する際に、入力された複数の入力テキストや入力情報中に同様の表現を行う単語が出現している場合に、該単語の重要度を下げる、または、格情報、品詞情報の出現を抑止する。

【0013】本発明（請求項6）は、要約語を選択する際に、常に案内する情報や、ユーザからの入力情報と重複する単語を含まないような重要語を選択すると共に、他のテキストと違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する。

10

20

30

40

50

【0014】本発明（請求項7）は、要約語を選択する際に、複数の入力テキスト間において他のテキストと同じ属性の要約語を選択し、該要約語に属性を付与し、該属性の共通部分を探索し、同じ表現を纏めて要約文を生成することが可能な要約語を選択し、選択された要約語を用いて要約文を生成する。

【0015】本発明（請求項8）は、要約語を選択する際に、ユーザから入力された入力情報から該ユーザの意図をユーザ意図推論知識を用いて推論し、該推論により抽出された単語に関連する該入力情報内の単語の重要度を上げる。本発明（請求項9）は、予め画面上に表示可能な文字数や行数を含む表示領域情報を保持しておき、要約文を生成した後に、該要約文と表示領域情報を比較し、要約文が表示領域情報に示される領域に対応するように、該要約文を編集する。

【0016】本発明（請求項10）は、要約文を編集する際に、要約文が表示領域情報に示される所定の範囲を越えないように文字を編集する、または、格情報に対応した表示情報を選択する。図2は、本発明の原理構成図である。本発明（請求項11）は、テキスト処理における要約文を生成する要約文生成装置であって、タイトル、本文、表を含む入力テキストや、ユーザからの入力情報を入力する文章入力手段10と、文章入力手段10で入力された入力テキストや入力情報の構造を解析する文章構造解析手段20と、出現位置と単語の係り受け関係で入力テキストや入力情報の単語の重要度を定義した重要単語定義テーブル101と、重要単語定義テーブル101を用いて単語の出現場所、重要度、重要語、該重要語の属性を取り出す重要語抽出手段30と、重要語抽出手段30で抽出した重要語に対し、出現場所と重要度で重み付けをして要約文に用いる要約語を選択する要約語選択手段40と、選択された要約語の属性に応じて生成する文のテンプレートを選択するテンプレート選択手段50と、選択されたテンプレートに要約語を埋め込んで要約文を生成する文生成手段60とを有する。

【0017】本発明（請求項12）は、要約語選択手段40において、常に案内する情報や、ユーザからの入力情報と重複する単語を含まないような重要語を選択する情報抑制手段を有する。本発明（請求項13）は、情報抑制手段は、常に案内する情報や、ユーザの入力情報と重複する単語を含む単語の重要度を下げる、または、フィルタリングによりカットする手段を含む。

【0018】本発明（請求項14）は、要約語選択手段40において、入力された入力テキスト及び入力情報について、他のテキストと違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較手段を有する。

【0019】本発明（請求項15）は、要約語比較手段

において、入力された複数の入力テキストや入力情報中に同様の表現を行う単語が出現している場合に、該単語の重要度を下げる、または、格情報、品詞情報の出現を抑止する手段を含。本発明（請求項16）は、要約語選択手段40において、常に案内する情報や、ユーザからの入力情報と重複する単語を含まないような重要語を選択する情報抑制手段と、他の文章と違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較手段とを有する。

【0020】本発明（請求項17）は、要約語選択手段40において、入力された複数の入力テキスト間において他のテキストと同じ属性の要約語を選択し、該要約語に属性を付与し、該属性の共通部分を探索し、同じ表現を纏めて要約文を生成することが可能な要約語を選択する手段を含み、要約文生成手段60は、選択された要約語を用いて要約文を生成する手段を含む。

【0021】本発明（請求項18）は、要約語選択手段40において、ユーザから入力された入力情報から該ユーザの意図をユーザ意図推論知識を用いて推論し、該推論により抽出された単語に関連する該入力情報内の単語の重要度を上げる手段を含む。本発明（請求項19）は、画面上に表示可能な文字数や行数を含む表示領域情報を更に有し、要約文生成手段60において、要約文を生成した後に、該要約文と表示領域情報を比較し、該要約文が表示領域情報に示される領域に対応するように、該要約文を編集する編集手段を含む。

【0022】本発明（請求項20）は、編集手段において、要約文が表示領域情報に示される所定の範囲を越えないように文字を編集する、または、格情報に対応した表示情報を選択する手段を含む。本発明（請求項21）は、テキスト処理における要約文を生成する要約文生成プログラムを格納した記憶媒体であって、タイトル、本文、表を含む入力テキストや、ユーザからの入力情報を入力させる文章入力プロセスと、文章入力プロセスにより入力された入力テキストや入力情報の構造を解析する文章構造解析プロセスと、出現位置と単語の係り受け関係で入力テキストや入力情報の単語の重要度を定義した重要単語定義テーブルを用いて単語の出現場所、重要度、重要語、該重要語の属性を取り出す重要語抽出プロセスと、重要語抽出プロセスで抽出した重要語に対し、出現場所と重要度で重み付けをして要約文に用いる要約語を選択する要約語選択プロセスと、選択された要約語の属性に応じて生成する文のテンプレートを選択するテンプレート選択プロセスと、選択されたテンプレートに要約語を埋め込んで要約文を生成する要約文生成プロセスとを有する。

【0023】本発明（請求項22）は、要約語選択プロセスにおいて、常に案内する情報や、ユーザからの入力

情報と重複する単語を含まないような重要語を選択する情報抑制プロセスを有する。本発明（請求項23）は、情報抑制プロセスにおいて、常に案内する情報や、ユーザからの入力情報と重複する単語を含む単語の重要度を下げる、または、フィルタリングによりカットするプロセスを含む。

【0024】本発明（請求項24）は、要約語選択プロセスにおいて、入力された入力テキスト及び入力情報について他のテキストと違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較プロセスを有する。

【0025】本発明（請求項25）は、要約語比較プロセスにおいて、入力された複数の入力テキストや入力情報中に同様の表現を行う単語が出現している場合に、該単語の重要度を下げる、または、格情報、品詞情報の出現を抑止するプロセスを含む。本発明（請求項26）は、要約語選択プロセスにおいて、常に案内する情報や、ユーザからの入力情報と重複する単語を含まないような重要語を選択する情報抑制プロセスと、他のテキストと違いのある属性の要約語を選択し、該要約語に属性を付与し、属性の共通部分または、差異部分を探索し、差異のある部分についてその差分を強調するような要約文を生成するための要約語を選択する要約語比較プロセスとを有する。

【0026】本発明（請求項27）は、要約語選択プロセスにおいて、入力された複数の入力テキスト間において他のテキストと同じ属性の要約語を選択し、該要約語に属性を付与し、該属性の共通部分を探索し、同じ表現を纏めて要約文を生成することが可能な要約語を選択するプロセスを含み、要約文生成プロセスにおいて、選択された要約語を用いて要約文を生成するプロセスを含む。

【0027】本発明（請求項28）は、要約語選択プロセスにおいて、ユーザから入力された入力情報から該ユーザの意図をユーザ意図推論知識を用いて推論し、該推論により抽出された単語に関連する該入力情報内の単語の重要度を上げるプロセスを含む。本発明（請求項29）は、要約文生成プロセスにおいて、要約文を生成した後に、該要約文と画面上に表示可能な文字数や行数を含む表示領域情報とを比較し、該要約文が該表示領域情報に示される領域に対応するように、該要約文を編集する編集プロセスを含む。

【0028】本発明（請求項30）は、編集プロセスにおいて、要約文が表示領域情報に示される所定の範囲を越えないように文字数や行数を編集する、または、格情報に対応した表示情報を選択するプロセスを含む。上記のように、本発明では、テキスト構造を解析し、テキストの種類を決定し、抽出したテキストの種類毎に係り受

け関係をもとにテキスト中に出現する単語の属性を決め、テキストの種類と単語の属性で重要度を定める。その後、テキストに対して重要度に応じて重要語を抽出し、抽出された重要語に対応するテンプレートを選択し、要約文を作成する。このように、本発明は、テキストの構造解析により単語の出現位置を考慮し、テキストの種類を解析しているため、タイトルや本文の先頭に出現する単語は重要であるという指定が可能となる。また、リンク情報は無視するという指定も可能である。

【0029】また、重要度を定める際に、案内対象が複数あった場合には、それぞれのテキストに対して抽出した重要語に対して、同様の表現を行う単語の重要度を下げ、ユーザの入力情報を解析し、ユーザの入力情報の表現中に存在した表現と同様の表現を行う重要度を下げる処理や、タイトル等、デフォルトでテキストの内容を解析し、デフォルトのテキスト内容中に存在した表現と同様の表現を行う単語の重要度を下げる処理を施すことで、タイトルと重複する項目、検索条件に指定した項目、他のテキストと共通の項目は要約文から削除できるので、有用な言葉のみで要約文を作ることが可能となる。

【0030】つまり、検索されたすべてのテキスト内容を比較し、それぞれのテキストの差分を強調し、例えば、「寝たきり老人や痴呆の老人を世話します」、「1人暮らし老人や低所得者層を世話します」と差分を強調し、表示することによって、ユーザのサービスの選択性を向上させることが可能となる。また、他の複数の文章と共通の表現があった場合に、それらの複数の文章に対する一つの要約文を集約して生成することが可能となる。

【0031】また、ユーザの入力情報に対応して、表示情報を変化させることが可能となり、ユーザに案内することが有効であろう情報を含んだ要約文を作成することが可能となる。例えば、「安い老人ホーム」で検索したことによって、ユーザがサービスの料金に興味を持っていることが分かった場合、要約文に料金情報を入れることによってユーザが欲しい要約文を作成することが可能となる。

【0032】また、装置毎に予め決められている表示領域に基づいて動的に表示する要約文の大きさを変化させることができるため、ユーザに対して的確な情報を表示することが可能となる。例えば、40文字という案内領域を考え、要約文の長さが長ければ単語の集約、要約文の長さが短ければ当該要約文の表示情報の増加を行うことで、ユーザに対して、的確な情報を表示することが可能となる。

【0033】

【発明の実施の形態】図3は、本発明の要約文作成装置の構成を示す。同図に示す要約文作成装置は、入力部10、文書構造解析部20、サービス重要語抽出部30、

要約語選択制御部40、テンプレート選択部50、要約文生成部60、出力部70、構造定義101、抽出ルール102、選択ルール103、シソーラス知識104、テンプレート選択ルール105から構成される。

【0034】入力部10は、テキスト及びユーザからの入力情報を入力する。本発明では、入力されるテキストとして、タグ情報を含むコーパスを例として説明するが、知識を設定することにより、文の構造情報（配置、フォント、等）により、本文等の種々の分類が可能であり、そのために必要となるテキストの配置や、フォント等を解析する処理を行えばよい。

【0035】文書構造解析部20は、テキストのタイトル、本文、表等出現位置やタグ情報等によりテキストの種別を定義する構造定義101に基づいて、テキストの種別・構造を解析する。詳しくは、入力されたテキストにおいて、解析対象となる文章の出現位置を考慮し、当該出現位置として、各種別毎に何番目の文章であるかの通し番号を付与する。これは、新聞記事等の場合、先頭に近い文について後に出現する文章とことなる重要度を設定したい場合等に使用することが可能となる。

【0036】サービス重要語抽出部30は、文書構造解析部20で抽出されたテキストの構造毎に形態素解析を行い、係り受け関係と品詞情報によりテキスト中に出現する単語の確信度を定義する抽出ルール102に従って、テキスト中の単語に対して、単語の属性と確信度、係り受け情報を設定する。本発明では、以下の方法を採用するものとする。

【0037】① 入力テキストに対して形態素解析を行うと共に、係り受け関係を解析する。

② テキストの係り受け関係と抽出ルール102を照合し、照合したものについて確信度を設定する。照合の方法としては、例としてパターンマッチン等がある。

【0038】要約選択制御部40は、サービス重要語抽出部30で抽出された単語に対して、テキストの種別と単語の属性、確信度より重要度を定義した選択ルール103と、単語の概念を定義したシソーラス知識104により近い概念の情報を統合し、重要度を算出し、重要度と係り受け関係より重要単語を決定することを制御する。テンプレート選択部50は、要約選択制御部40により抽出された重要単語に基づいて、重要単語群に対して定義されているテンプレート選択ルール105を参照して、要約文を作成するためのテンプレートを選択する。

【0039】要約文生成部60は、テンプレート選択部50で選択されたテンプレートに重要単語を当てはめることによって要約文を作成する。出力部70は、作成された要約文を出力する。ここで、要約語選択制御部40の構成について説明する。図4は、本発明の要約語選択制御部の構成（その1）を示す。

【0040】同図に示す要約語選択制御部40は、要約

語選択部 4 1 及び情報抑制部 4 2 から構成される。要約語選択部 4 1 は、サービス重要語抽出部 3 0 で抽出された単語に対して、テキストの種別と単語の属性、確信度より重要度を定義した選択ルール 1 0 3 と、単語の概念を定義したシソーラス知識 1 0 4 により近い概念の情報を統合し、重要度を算出し、重要度と係り受け関係より重要単語を決定し、情報抑制部 4 2 に転送する。

【0041】情報抑制部 4 2 は、形態素解析等によってユーザの入力を解析した結果やデフォルトで案内される案内内容を解析した結果を記憶する抑制語記憶部 1 0 6 を参照して、重要単語について、ユーザの入力表現中に存在した表現とデフォルトで案内される表現と同様の表現を行う単語の重要度の値を下げたり、フィルタリングによりカットする処理を行う。

【0042】また、要約語選択制御部 4 0 の他の構成について説明する。図 5 は、本発明の要約語選択制御部の構成（その 2）を示す。同図に示す要約語選択制御部 4 0 は、要約語選択部 4 1 と要約語比較部 4 3 から構成される。要約語選択部 4 1 は、決定された重要単語について、テキストが複数存在する場合に、各テキストに対して行った一連の処理結果をテキスト毎に記憶しておく要約語記憶部 1 0 7 を参照して、複数のテキスト中に同様の表現を行う単語が出現していた場合に、要約語比較部 4 3 に転送する。

【0043】これにより要約語比較部 4 3 は、同様の単語の属性を比較し、タイトル等、デフォルトで案内される内容を解析し、デフォルトの案内内容中に存在する表現と同様の表現を行う重要単語の重要度を下げる、または、格情報、品詞情報の出現を抑止する。また、要約語選択部 4 1 において、複数のテキスト中に同様の表現を行う単語を取得した場合に、上記の例では、要約語比較部 4 3 に転送しているが、これをテンプレート選択部 5 0 に転送し、前述の処理を施してテンプレートを取得し、要約文生成部 6 0 において、複数のテキストに関する 1 つの要約文を生成することも可能である。つまり、類似する複数のテキストの要約文を作成する。

【0044】次に、ユーザの意図を推論し、ユーザが欲している情報に関する表現の重要度上げる場合について説明する。図 6 は、本発明の要約語選択制御部の構成（その 3）を示す。同図に示す要約語選択制御部 4 0 は、要約語選択部 4 1、ユーザ意図判定部 4 4、選択ルール 1 0 3、シソーラス知識 1 0 4、及びユーザ意図推論知識 1 0 8 から構成される。

【0045】ユーザ意図判定部 4 4 は、ユーザから入力された情報内に含まれる重要単語について、ユーザ意図推論知識 1 0 8 を参照して当該重要単語に対応するユーザが意図しているであろう単語を抽出し、要約語選択部 4 1 に転送する。これにより、要約語選択部 4 1 は、ユーザ意図判定部 4 4 において抽出された単語に重要度を付与する。このとき、複数の重要単語があった場合に、重

要度の最も高いものを重要単語とする。

【0046】次に、表示画面において所定の案内領域が決定されており、当該領域に要約文が収まるように編集する場合について説明する。図 7 は、本発明の要約文生成部の構成を示す。要約文生成部 6 0 において生成された要約文を要約文編集部 6 5 に転送する。要約文編集部 6 5 において、表示領域情報 1 0 9 を参照し、当該案内領域に示されている表示部の広さ、表示情報量、表示情報属性と要約文を比較して、要約文が案内領域に示される情報量を越える場合には、当該要約文を単語の集約を行い、要約文が案内領域に示される情報量の所定の閾値より小さければ表示情報（要約文）を増加させる等の処理を行う。

【0047】

【実施例】以下、図面と共に本発明の実施例を説明する。以下、図 3～図 5 の構成に基づいて説明する。まず、図 3 に示す各構成要素で参照するルール等について説明する。図 8 は、本発明の一実施例の構造定義の例を示す。

【0048】同図に示す構造定義 1 0 1 は、単語の出現場所、重要度、重要語、該重要語の属性を取り出すもので、文章構造解析部 2 0 により参照される。適用文書形態、ルール及び文種別からなり、ルールから選択される文種別は、文章の構造の種別である、表、図、メニュー、等以外に記入されている内容の種別である。本文、タイトル、手続、連絡先、リンク等も用いることもできる。

【0049】図 9 は、本発明の一実施例の抽出ルールの例を示す。同図に示す抽出ルール 1 0 2 は、係り受け関係と品詞情報により文章中に出現する単語の確信度を定義したものであり、サービス重要語抽出部 3 0 により参照され、係り受け関係、属性及び、確信度から構成される。当該抽出ルール 1 0 2 の記述の例として、特願平 1 0-53889 の形態を用いることも可能である。この方法は、同じ代表分類ノードに属する単語は同じものとして扱うことにより、短い文章でも頻度に基づいてキーワードの重要度を判断することができる。

【0050】また、格文法に基づいて、動詞の場所格、頻度等を用いることも可能である。さらに、確信度 0 を定義し、抽出しないというルールを記述することができる。図 1 0 は、本発明の一実施例の選択ルールの例を示す。同図に示す選択ルール 1 0 3 は、文章の種別と単語の属性、確信度より重要度が定義されており、要約語選択制御部 4 0 により参照される。選択ルールと重要度からなり、重要度は、確信度に対して四則演算を行い、その結果を統合し、単語毎の重要度として算出したものである。

【0051】図 1 1 は、本発明の一実施例のシソーラス知識の例を示す。同図に示すシソーラス知識 1 0 4 は、単語の概念が定義されており、要約語選択制御部 4 0 に

より参照される。単語の意味、上下関係の他に、その単語自体の抽象度、具象度が定義されている。動詞の場合は、格の情報が定義されている。図12は、本発明の一実施例のテンプレートの例を示す。

【0052】同図に示すテンプレート105は、要約語を当てはめるものでテンプレート選択部50により選択され、要約文生成部60で使用される。解析項目、抽出された重要語、条件、及びこれらに対応する複数のテンプレート群から構成される。以下に本実施例の動作を説明する。

【第1の実施例】図13、図14は、本発明の第1の実施例のHTML文書の例であり、一連のHTML文書である。以下では、このHTML文書を入力とした場合の動作を説明する。

【0053】図15は、本発明の第1の実施例の動作を示すフローチャートである。ステップ101) 文章構造解析部20は、入力部10から入力されたHTML文書のタグ情報、文書情報に対し、構造定義101を適用して、文書の文毎に文番号を振り、文種別を割り振る。文番号は文単位、または、表の場合は、表の属性単位で分けるものとする。商品が並列で並んでいる場合には、商品単位で分けるものとする。入力された文書の解析結果を図16に示す。同図の例において、文種別は「メニュー」「タイトル」「本文」「表」「手続き」「連絡先」「リンク」に分けられ、各々の文種別毎に文に文番号が付与されている。

【0054】ステップ102) 次に、サービス重要語抽出部30において、それぞれの文種別の文番号毎に形態素解析を行い、その後、例えば、特開平10-2669827に開示されている方法により情報を加工し、その情報に基づいて動詞の格情報を抽出ルール102に基づいて取り出す。なお、特開平10-2669827に開示されている方法は、問い合わせ文の形態素・構文解析を行い、単語単位に分割し、当該問い合わせ文中に職業名情報が含まれている場合、知識ベースの変換テーブルを参照して、職業語尾を動詞に変換し、職業名情報の語尾以外の部分は商品名とし、当該商品名と動詞より関連知識を用いて職業分類名を推論し、推論した結果でデータベースを検索する。一方、問い合わせ文中に商品名と動詞が含まれている場合は、変換テーブルの参照を行わずに、直接商品名と動詞より関連知識を用いて職業分類名を推論しデータベースを検索する方法である。

【0055】解析結果を纏める方法は種々あるが、本実施例では、「商品と商品に対応する動詞と動詞に対応する対象者について」、「対象者について」、「手続きについて」、「連絡先について」の4つの項目について解析を行った。並列併記される情報の場合、それぞれに対して確信度が与えられる。サービス重要語抽出部30により抽出された例を図17に示す。

【0056】ステップ103) 次に、要約語選択制御

部40は、サービス重要語抽出部30から取得した図17に示す文種別と確信度より、選択ルール103を用いて重要度を定義する。要約語選択制御部40により重要度の定義の例を図18に示す。ステップ104) さらに、要約語選択制御部40において、シソーラス知識104を用いて、概念の統一、表記が同じレコードをマージ、共通概念を一つに纏める処理を行う。この時、「等」などで省略されている表示を補完する。

【0057】この結果を図19に示す。ステップ105) さらに、要約語選択制御部40において、それぞれ解析した項目について重要度の高いレコードを選択する。選択された結果を図20に示す。ステップ106)

要約語選択制御部40において、形態素解析の結果得られる動詞の格構造に基づいて情報を補完する。本実施例では、「商品と商品に対応する動詞、動詞に対応する対象者について」の動詞、『給付・助成』は、対象格に品詞が名詞、意味が人の情報をとることがわかるため、「対象者について」の対象者を利用することで情報を補う。「手続きについて」の動詞、『提出』は対象格に品詞が名詞、意味が人の情報を取らないことがわかるため、情報は補完されない。

【0058】補完された例を図21に示す。ステップ107) 次に、テンプレート選択部50が、選択された単語に基づいてテンプレート105を参照し、適切なテンプレートを選択する。この例では、「商品と商品に対応する動詞、動詞に対応する対象者について」は、『「対象者」に「商品」を「動詞」します。』というテンプレートを選択し、「手続きについて」は、『「場所格」に「商品」を「動詞」してください。』というテンプレートを選択し、「連絡先について」は、『詳しくは、「連絡先部署課名」(電話「電話番号」)まで。』というテンプレートを選択する。

【0059】ステップ108) 次に、要約文生成部60は、テンプレートに当てはめる要素を加工する。シソーラス知識104を用いて、品詞の一般性を考え、抽象の表現が使われていた場合、可能なら具象の表現へ置き換える。なお、項目数が多い場合には、後半分をカットし、『等』という表現へと置き換える。この場合、「商品と商品に対応する動詞、動詞に対応する対象者について」内の情報『日常生活用具』が、“ガス漏れ警報機、自動火災感知機、移動感知機、自動消化器、マットレス、体位変換器”へと変換され、さらに、“ガス漏れ警報機、自動火災感知器、移動感知機 等”への変換される。また、“寝たきり高齢者、一人暮らし高齢者、痴呆性高齢者”が、“寝たきり、一人暮らし、痴呆性高齢者”へと変換される。

【0060】ステップ109) 要約文生成部60において、出力する要約文を選択する。この場合は、重要度により、『商品と商品に対応する動詞、動詞に対応する

対象者について』が選択される。ステップ110) 出力部70において、要約文を出力する。この例の場合は、『ガス漏れ警報機、自動火災感知機、移動感知機、等を寝たきり、一人暮らし、痴呆性高齢者に給付・助成します』が出力される。

【0061】[第2の実施例] 次に、本実施例では、前述の第1の実施例のフローチャートにおけるステップ103において、必ずユーザに案内する情報やユーザの入力とを比較し、重要度を下げる場合について、図4の構成に基づいて説明する。例えば、タイトルを必ず表示する場合、

『タイトル：ホームヘルパーの派遣

本文：身体上または、精神上の障害のため、日常生活を営むのに支障のある概ね65歳以上の人がいる家庭にホームヘルパーを派遣します。

【0062】主なサービスは、洗濯・掃除・買物・食事の支度・身の回りのお世話・相談などです。但し、生計中心者の前年所得額により一部費用を負担して頂きます。』について、形態素解析の結果を利用した結果、タイトルから、

“商品：ホームヘルパー（人） 対応動詞：派遣 重要度：9”、

本文から、

“商品：ホームヘルパー（人） 対応動詞：派遣 重要度：6”、

商品：洗濯・掃除・買物・食事の支度・身の回りのお世話・相談など（事）重要度：6、

商品：費用（物） 動詞：負担 重要度1”

となる。商品と動詞の組み合わせでタイトルと全く同じ物が抽出された場合、それらの情報の重要度を0と設定し、その後、ステップ104のマージへと進む。その後、処理を進め、テンプレート選択ルールとして、『商品だけで動詞がなく、商品の意味が「事」だったら、「商品（事）」を行います』というルールを適用すると、結果として、『洗濯・掃除・買物等を行います』という要約文が得られる。

【0063】本実施例は、タイトルを必ずしも表示する場合、タイトルにあった単語の重要度を下げているが、ユーザの入力についても同様に、ユーザの入力にあった単語の重要度を下げることもできる。

【第3の実施例】本実施例では、前述の第1の実施例のステップ103において、複数のサービスの情報を比較し、他と違った表現の重要度を上げる例を図5に基づいて説明する。

【0064】入力文として図22に示す入力文を入力して、ステップ101、ステップ102の解析処理により、本文1から図23(a)に示す内容が、また、本文2から図23(b)に示す内容が抽出される。ここで、図23(a)の①と図23(b)の④、図23(a)の③と図23(b)の⑥が全く同じ表現のため、重要度を

大きく下げる。この例では、“-4”とする。また、図23(a)の②と図23(b)⑤はレコード内容のうち、1つのレコードのみが違っている表現のため、重要度を下げる。この例では、“-2”とする。

【0065】その結果、重要度はそれぞれ図24に示すように設定され、その後、本文1、本文2のそれぞれに情報をもとにステップ104のマージへと進む。その後、処理を進め、テンプレート選択ルールとして、動詞が「預かる」で、商品に「人」の意味があり、「場所格」、「時間格」がある場合は、『「対象者」を「時間」に「場所」で預かります』というルールを適用すると、結果として、本文1からは、『痴呆性高齢者を夜間に老人ホームで預かります』本文2からは、『痴呆性高齢者を日中に老人ホームで預かります』が得られる。

【0066】[第4の実施例] 前述の第3の実施例のステップ103において、複数のサービスの情報を比較し、他と同じ表現の重要度を下げると共に、対応しているレコードに対して格情報、品詞情報の出現を抑止する例を図4及び図5に基づいて説明する。前述の第3の実施例と同様に、図22に示す入力を行った場合、ステップ101、ステップ102の処理により、図25の結果が抽出される。

【0067】ここで、本文1と本文2で、「動詞」と「商品」が一致しているものを比較し、内容が全部同一ならそのレコードの重要度を「0」として棄却し、内容が一部でも違っているなら、動詞情報を残し、違っている格情報のみを残し、他の格情報や品詞情報を捨てる。その結果、本文1において、

②動詞：預かる 時間：日中 重要度：4
を設定し、

本文2において、

⑤動詞：預かる 時間：夜間 重要度：4

を設定する。その後、本文1、本文2のそれぞれの情報をもとにステップ104のマージの処理に以降する。

【0068】その後、処理を進め、テンプレート選択ルールとして、『「商品」がなく、「動詞」が預かるで、「時間格」がある場合は、「時間」に預かります。』というルールを適用すると、結果として、本文1からは、『夜間預かります』本文2からは、『日中預かります』という要約文が得られる。

【0069】[第5の実施例] 本実施例は、図5に示す構成に基づいて、要約語選択部41において、複数のテキスト中に同様の表現を行う単語を取得した場合に、要約文生成部60において、複数のテキストに関する1つの要約文を生成する例を説明する。つまり、複数のサービスの情報を比較し、他と同じ表現を纏めることによって、複数のテキストから1つの要約文を生成するものである。

【0070】図5において、要約語選択部41におい

て、各テキストに対して行った一連の処理結果をテキスト毎に記憶しておく要約語記憶部107を参照し、以下の内容を取得したものとす。要約語記憶部107に格納されている本文1からは、

①商品：介護（事） 動詞：困難 対象者：痴呆性高齢者 重要度：6

②商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 時間：日中重要度：6

③商品：負担 動詞：軽減 対象者：介護者 重要度：6

また、本文2からは、

①商品：介護（事） 動詞：困難 対象者：痴呆性高齢者 重要度：6

②商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 時間：夜間重要度：6

③商品：負担 動詞：軽減 対象者：介護者 重要度：6

が抽出されたものとする。

【0071】ここで、本文1と本文2で、動詞と商品が一致しているものを要約語比較部43で比較し、内容が全部同一なら、要約語選択部41においてそのレコード内容を複数の文章のまとめの文として採用するように、以下の内容をテンプレート選択部50に通知する。

商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム

テンプレート選択部50では、テンプレート選択ルール105を参照して、『商品が“人”、動詞が“預かる”、場所格がある場合には、「商品（人）」を「場所」で預かります。』というルールを適用することとする。

【0072】これを要約文作成部60において、当該テンプレートを使用して、『痴呆性高齢者を老人ホームで預かります』という1つの要約文を生成する。

〔第6の実施例〕本実施例では、ユーザから入力された入力情報からユーザの意図を推論して、当該ユーザの欲している情報に関する表現の重要度を上げる例を、図6を用いて説明する。

【0073】要約語選択制御部40において、重要語抽出部30で抽出された重要語を取得すると、当該重要語に基づいてユーザ意図判定部44において、ユーザ意図推論知識108を検索する。本実施例では、ユーザから入力された入力情報が、『安くおじいちゃんを世話して欲しい』であるとき、データベースを検索すると、

『タイトル：日中介護サービス

本文：在宅の痴呆性高齢者を老人ホームでおかずかりします。

【0074】実施回数：サービスの利用回数は年30回まで。

提供時間：午前9時から午後5時。

利用者負担：1,050円/回』

が検索結果として検索され、それを構文解析部20と重要語抽出部30で処理することにより、

・商品と商品に対応する動詞、動詞に対する対象者について：

①商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 重要度：6

・商品の実施回数について：

②回数：年30回 重要度：2

・商品の時間について：

10 ③時間：午前9時～午後5時 重要度：2

・商品の値段について：

④値段：1050円/回 重要度：2

が抽出さえる。ここで、ユーザ意図判定部44において、ユーザからの入力情報「安くおじいちゃんを世話して欲しい」に基づいて意図推論知識108を参照する。この場合、入力情報中に「安く」というキーワードが入っていたため、ユーザについて値段情報を重視していることを判断し、値段情報についての重要度を計算する。

【0075】ここでは、ユーザからの入力情報に対応する情報の重要度を3倍として

・商品の値段について：

④値段：1050円/回 重要度：6

となる。その結果、要約語選択部41では、

①商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 重要度：6

②回数：年30回 重要度：2

③時間：午前9時～午後5時 重要度：2

④値段：1050円/回 重要度：6

30 が抽出される。これらの情報は、次段のテンプレート選択部50に転送される。

【0076】テンプレート選択部50では、テンプレート選択ルール105として、動詞が“預かる”で、商品に“人”の意味があり、“場所格”、“時間格”があり、“回数”、“値段”、“時間”の情報が有る場合には、『「対象者」を「回数」、「時間」の間、「値段」で「場所」にて預かります。』というルールを適用する。これにより、要約文生成部60では、当該テンプレートを利用して、『「痴呆性高齢者」を、「年30回」、「午前9時～午後5時」の間、「1050円/回」で「老人ホーム」にて預かります。』という要約文を生成し、出力する。

【0077】〔第7の実施例〕次に、本実施例では、図7に基づいて、要約文生成部60において生成された要約文を所定の表示領域情報109に基づいて編集する例を説明する。まず、要約文生成部60で要約候補文として、『痴呆性高齢者を、年30回、午前9時～午後5時の間、1050円/回で老人ホームにて預かります。』が作成されたとする。

50 【0078】ここで、表示領域情報109を参照すると、全文字数が46文字であるため、「文字数30文字

以内」の条件をみたしていないとし、上記の要約候補文を加工する。ここで、要約文編集部65は、テンプレート選択部50で選択された、『「対象者」を「回数」、「時間」の間、「値段」で「場所」にて預かります。』というテンプレート内のそれぞれの要素の重要度を参照し、

「対象者」：痴呆性高齢者	重要度：6
「回数」：年30回	重要度：2
「時間」：午前9時～午後5時	重要度：2
「値段」：1050円/回	重要度：6
「場所」：老人ホーム	重要度：6

から、重要度の低い「回数」、「時間」のうち、文字数の多い「時間」を消去する。その結果、

「対象者」：痴呆性高齢者	重要度：6
「回数」：年30回	重要度：2
「値段」：1050円/回	重要度：6
「場所」：老人ホーム	重要度：6

となり、ここで、テンプレート選択部50でテンプレートを再度選択し、動詞が“預かる”で、商品に“人”の意味があり、場所格、時間格があり、回数、値段の情報がある場合には、『「対象者」を「回数」、「値段」で「場所」にて預かります。』というテンプレートを選択し、要約文生成部60において、再度、「痴呆性高齢者を、年30回、1050円/回で老人ホームにて預かります。」の34文字の要約文候補を作成する。これについても同様に、表示領域情報109を参照すると、「文字数30文字以内」の条件を満たしていないため、当該要約文候補を再度加工する。ここでは、上記の重要度に基づいて、「回数」を消去するものとする。この結果、「痴呆性高齢者を、1050円/回で老人ホームにて預

かります。」(読み：ちほうせいこうれいしゃをせんどじゅうえん/かいろうじんほむであずかります)を得る。この要約文候補は、となり、表示領域情報109の内容を満たしている。

【0079】上記の例は、文字数が多過ぎるため加工した例であるが、文字数が少ない場合には、複数の文を選択したり、前述の第1の実施例で加工した際に省略された情報を元に戻す等の処理が可能である。なお、文字数が少ないという判定は、表示領域情報109の決められている文字数の80%以下の場合、等の閾値を設定しておくことによりその判断が可能となる。

【0080】また、上記の実施例では、図15に示すフローチャートに基づいて説明してるが、同図のフローチャートをプログラムとして構築し、要約書作成装置として利用されるコンピュータに接続されるディスク装置や、フロッピーディスク、CD-ROM等の可搬記憶媒体に格納しておき、本発明を実施する際にインストールすることにより、容易に本発明を実現できる。

【0081】なお、本発明は、上記の実施例に限定され

ることなく、特許請求の範囲内で種々変更・応用が可能である。

【0082】

【発明の効果】上述のように、本発明によれば、単語の出現位置を考慮し、テキストの種類を解析することにより、タイトルや本文の先頭に出現する単語は重要であると認識する、タイトルと重複する項目、検査条件に指定した項目、他の案内文と共通の項目は要約文から削除可能となることから、有用な単語のみで要約文を作成することができる。

【0083】また、複数のテキスト間における差分を強調することが可能であるため、情報を取捨選択するために有効な要約文を提供することができる。また、ユーザからの入力情報に基づいて、ユーザの意向を推論することにより、ユーザにニーズに合った要約文を作成することができる。また、複数のテキスト間において共通の表現がある場合には、テキスト間の共通情報を抽出し、個々の要約文ではなく、代表する要約文を1つに纏めて作成することができる。

【0084】また、装置上の表示領域の制限や、表示時間の制限がある場合に、要約文を領域や時間に応じて、主要部分を欠落させることなく動的に変更することが可能となり、ユーザに対して的確な要約文を提示することができる。このように、本発明では、タイトル等、デフォルトで案内する情報と、ユーザの入力を解析し、要約文中に同様の表現を行う単語の出現を抑制するため、表示情報領域が限られたモバイル端末における情報案内や、案内時間が限られた音声による情報案内、検索結果が多く得られた際に、有効な情報を捜し出す検索エンジンやイエローページサービスにおける要約文の作成に有効である。

【0085】複数の案内情報等の文章の差分を強調し、表示するため、検索エンジンやイエローページサービスにおいて検索結果と共に表示することによって、ユーザの情報を選択肢を増すことができる。

【図面の簡単な説明】

【図1】本発明の原理を説明するための図である。

【図2】本発明の原理構成図である。

【図3】本発明の要約文作成装置の構成図である。

【図4】本発明の要約語選択制御部の構成図(その1)である。

【図5】本発明の要約語選択制御部の構成図(その2)である。

【図6】本発明の要約語選択制御部の構成図(その3)である。

【図7】本発明の要約文生成部の構成図である。

【図8】本発明の一実施例の構造定義の例である。

【図9】本発明の一実施例の抽出ルールの例である。

【図10】本発明の一実施例の選択ルールの例である。

【図11】本発明の一実施例のシソーラス知識の例であ

る。

【図12】本発明の一実施例のテンプレートの例である。

【図13】本発明の第1の実施例のHTML文書の例(その1)である。

【図14】本発明の第1の実施例のHTML文書の例(その2)である。

【図15】本発明の第1の実施例の動作を示すフローチャートである。

【図16】本発明の第1の実施例の構文解析結果の例である。

【図17】本発明の第1の実施例のサービス重要語抽出部により抽出された結果の例である。

【図18】本発明の第1の実施例の重要度設定の例である。

【図19】本発明の第1の実施例の要約語選択制御部による概念の統一・マージ・共通概念の統一処理の例である。

【図20】本発明の第1の実施例の要約語選択制御部による重要度の高いレコードの選択処理の例である。

【図21】本発明の第1の実施例の要約語選択制御部における情報補充の例である。

【図22】本発明の第3の実施例の入力文の例である。

【図23】本発明の第3の実施例の種別付与・重要語抽出結果である。

【図24】本発明の第3の実施例の重要度付与の例であ*

る。

【図25】本発明の第4の実施例の種別付与・重要語抽出された結果の例である。

【符号の説明】

- 10 入力部
- 20 文章構造解析手段、文章構造解析部
- 30 重要語抽出手段、サービス重要語抽出部
- 40 要約語選択手段、要約語選択制御部
- 41 要約語選択部
- 42 情報抑制部
- 43 要約語比較部
- 44 要約語選択部
- 50 テンプレート選択手段、テンプレート選択部
- 60 要約文生成手段、要約文生成部
- 65 要約文編集部
- 70 出力部
- 101 構造定義
- 102 抽出ルール
- 103 選択ルール
- 104 シソーラス知識
- 105 テンプレート選択ルール
- 106 抑制語記憶部
- 107 要約語記憶部
- 108 ユーザ意図推論知識
- 109 表示領域情報

【図1】

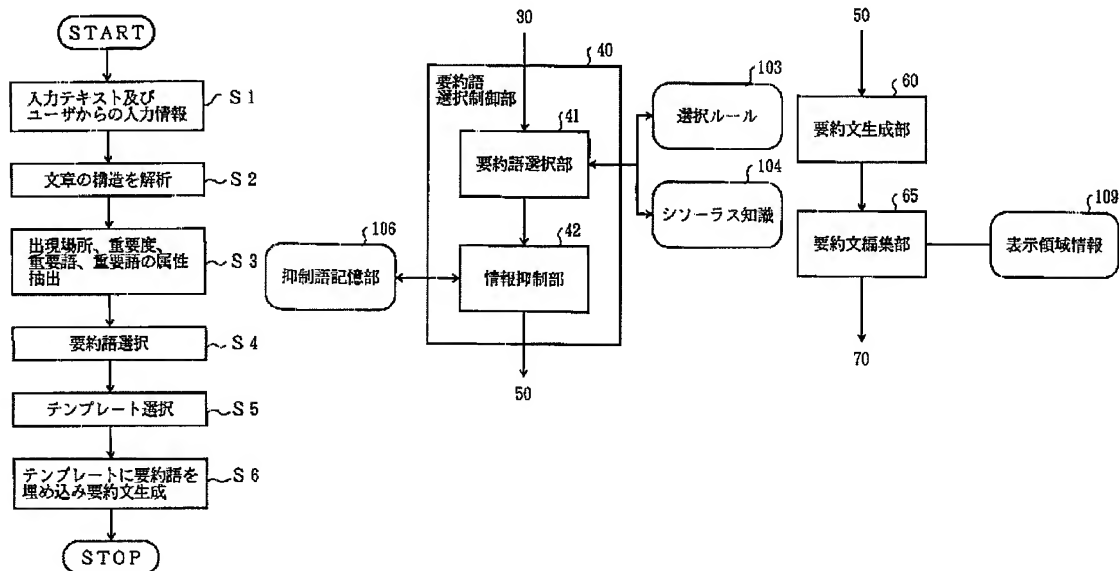
【図4】

【図7】

本発明の原理を説明するための図

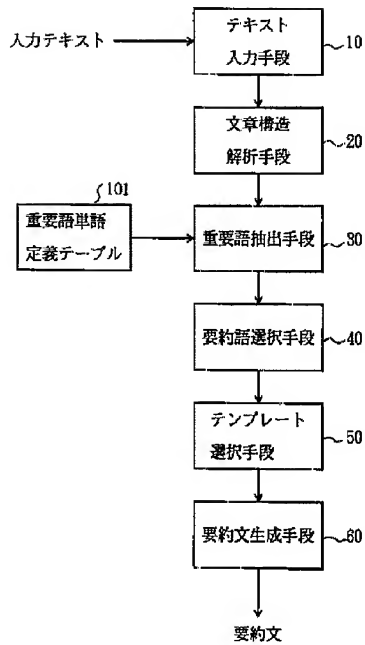
本発明の要約語選択制御部の構成図(その1)

本発明の要約文生成部の構成図



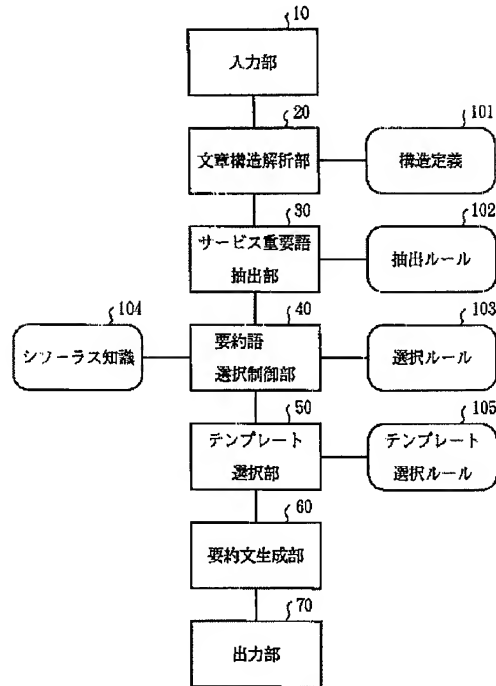
【図2】

本発明の原理構成図



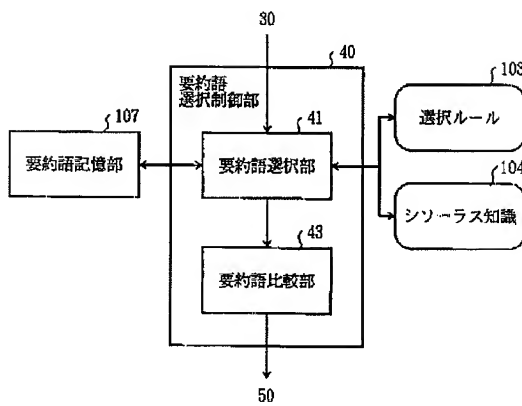
【図3】

本発明の要約文作成装置の構成図



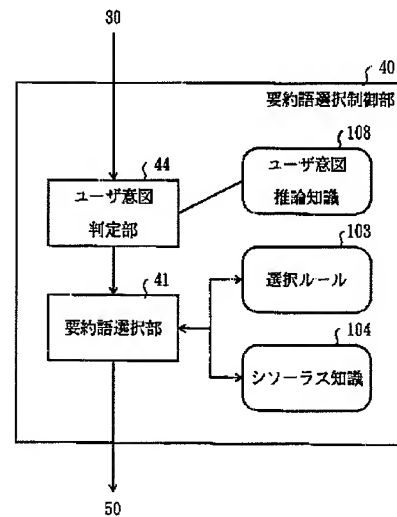
【図5】

本発明の要約語選択制御部の構成図（その2）



【図6】

本発明の要約語選択制御部の構成図（その3）



【図8】

本発明の一実施例の構造定義の例

101	
適用文書形態	ルール
HTML	・文書の先頭で、センタリングタグとフォントが強調（文字ポイント数大、色文字、点線、アンダーライン等の書式変更等）されている ・重要動詞の対象格に名詞がきており、その名詞が事の意味であるとき、その名詞 動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞 商品とベアになっている動詞 「サービス」の前に付いている複合名詞 重要動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞 動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞 物、事の意味を持つ名詞 対象者の意味を持つ名詞 重要動詞の対象格に名詞がきており、その名詞が人の意味であるとき、その名詞 動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞
	・表のタグの中に含まれている
	・画像のタグの中に含まれている
	・リンクのタグの中に含まれている
電子テキスト (メール等)	・センタリング・強調された「問い合わせ・相談は」という項目につな がっている部分 ・タイトルの次の行からの一連の文章で、表、図、リンク等の属性に当 てはまらない部分 ・複数のリンク情報が順次よく1段ずつ上位へと上がっていく範囲に記 述されている場合 ・センタリング・強調された「申請手続き」という項目につながる一連 の部分 ・文書の先頭で、●等の記号の後に記述 ・緊密に囲まれている ・引用の印(〇)が先頭、もしくは半角ローマ字の後に付いている ・文末でキャラクタリーラインで区切られている ・表、リンクネチャ以外

【図10】

本発明の一実施例の選択ルールの例

103	
選択ルール	重要度
「タイトル」中の「商品」	確信度の3倍
「図」中の情報	0
「問い合わせ先」内に出現した「商品」	0
「本文」中の「商品」	確信度の2倍
表示情報、ユーザの入力に出現した単語：確信度下げる	

【図9】

本発明の一実施例の抽出ルールの例

102	
属性	確信度
重要動詞の対象格に名詞がきており、その名詞が事の意味であるとき、その名詞	商品 3
動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞	商品 1
商品とベアになっている動詞	動詞 1
「サービス」の前に付いている複合名詞	商品 1
重要動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞	商品 3
動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞	商品 1
物、事の意味を持つ名詞	商品 1
対象者の意味を持つ名詞	対象者 1
重要動詞の対象格に名詞がきており、その名詞が人の意味であるとき、その名詞	動詞の対象格 1
動詞の対象格に名詞がきており、その名詞が物の意味であるとき、その名詞	動詞の対象格 1

【図11】

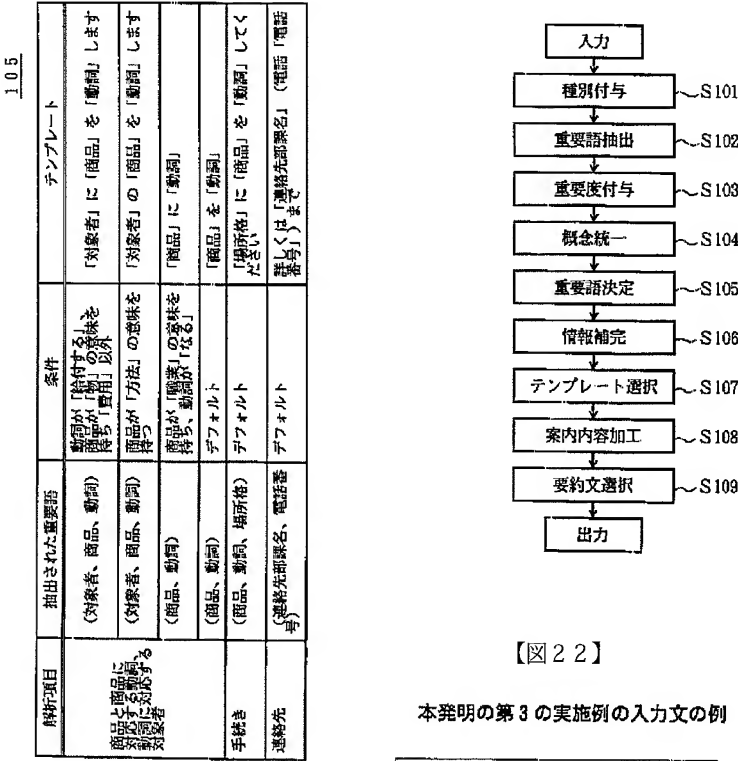
本発明の一実施例のシソーラス知識の例

104	
品詞：実体、上位、分類フラグ	
名詞：物、ルート、抽象	
名詞：日常生活用具、物、抽象	
名詞：自動消化器、日常生活用具、具体	
名詞：自動火災感知器、日常生活用具、具象	
名詞：ガス漏れ警報機、日常生活用具、具象	
名詞：マットレス、日常生活用具、具象	

【図12】

【図15】

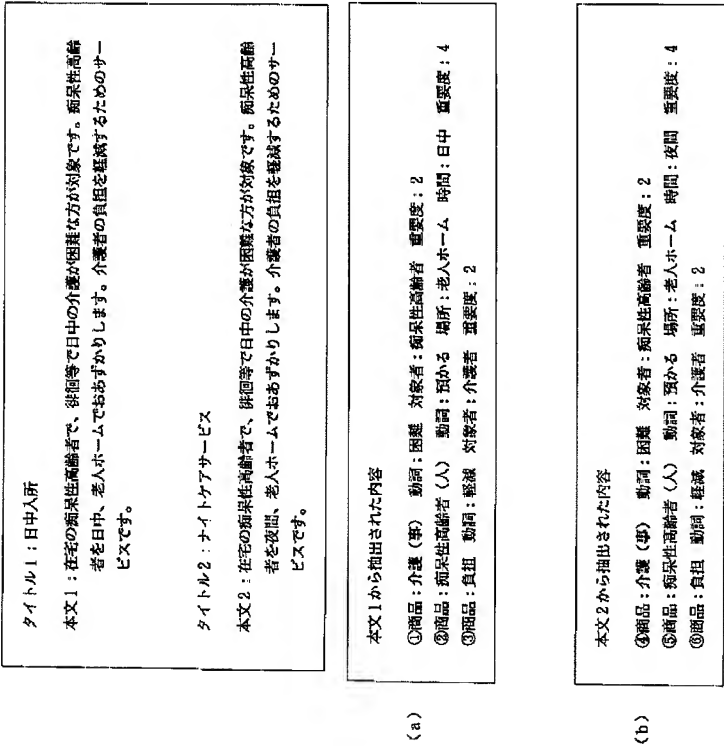
本発明の一実施例のテンプレート選択ルール例 本発明の第1の実施例の動作を示すフローチャート



【図22】

【図24】

本発明の第3の実施例の入力文の例 本発明の第3の実施例の重要度付与の例



【図13】

本発明の第1の実施例のHTML文書の例（その1）

```

<html>
<head>
<meta http-equiv="Content-Type"
content="text/html; charset=x-sjis">
</head>
  <A HREF="../gyousei.htm"><font size=3><I><行政サービス></I></font></A>
  <A HREF="../../fukusi.htm"><font size=3><I><福祉サービス></I></font></A>
<HR>
<CENTER>
<h1 align="center"> 日常生活用具の給付等</h1>
<TABLE BORDER="0" WIDTH="610">
<TR><TD><BLOCKQUOTE>
  <font size="4"> 家庭での介護を助けるサービスです。ねたきり高齢者、ひとり暮らし高
  齢者、痴呆性高齢者の介護のための日常生活用具を給付又は助成します。このサービスを受
  けるためには介護者に所得制限があります。
  </font></BLOCKQUOTE></TD></TR>
<TR><TD ALIGN="center">
  <font color="#DD0022" size="4"> 給      付</font>
</TD></TR>
<TR><TD ALIGN="center">
  <table border="1" CELLPADDING="3">
    <tr>
      <TH>ひとり暮らし高齢者</TH>
      <td>自動消化器、ガスもれ警報器</td>
    </tr>
  </table>
</TD></TR>
<TR><TD ALIGN="center">
  <font color="#DD0022" size="4"> 貸      与</font>
</TD></TR>
<TR><TD ALIGN="center">
  <table border="1" CELLPADDING="3">
    <tr>
      <TH>ひとり暮らし高齢者</TH>
      <td>自動火災感知機</td>
    </tr>
  </table>

```

【図14】

本発明の第1の実施例のHTML文書の例（その2）

```

<tr>
  <th>痴呆性高齢者</th>
  <td>移動感知機</td>
</tr>
</table>
</td></tr>
<tr><td align="center">
<font color="#DD0022" size="4"> 購入費の助成</font>
</td></tr>
<tr><td align="center">
<table border="1" cellpadding="3">
  <tr>
    <th>ねたきり高齢者</th>
    <td>マットレス、体位変換器</td>
  </tr>
</table>
<tr><td align="center">
<font color="#DD0022" size="4"> 申請手続き</font>
</td></tr>
<tr><td align="center">
申し込みは、老人課へ申請書を提出していただきます。<br>
</td></tr>
<tr><td align="center">
<font color="#DD0022" size="4"> 問い合わせ・相談は</font>
</td></tr>
<tr><td align="center">
老人課      ( 9 9 ) 9 9 9 9<br>
</td></tr>
</table>
<hr>
<br>
<a href="../gyousei.htm"> 行政サービス</a><br>
</center>
</body>
</html>

```

【図16】

本発明の第1の実施例の構文解析結果の例

文種別：メニュー

- [1] 行政サービス
- [2] 福祉サービス

文種別：タイトル

- [1] 日常生活用具の給付等

文種別：本文

- [1] 家庭での介護を助けるサービスです。
- [2] ねたきり高齢者、ひとり暮らし高齢者、痴呆性高齢者の日常生活用具や介護用具を給付又は助成します。
- [3] このサービスを受けるためには介護者に所得制限があります。

文種別：表

- [1] 給付
- [2] ひとり暮らし高齢者
- [3] 自動消化器
- [4] ガスもれ警報器
- [5] ひとり暮らし高齢者
- [6] 自動火災感知機
- [7] 痴呆性高齢者
- [8] 移動感知機
- [9] 購入費の助成
- [10] ねたきり高齢者
- [11] マットレス
- [12] 体位変換器

文種別：手続き

- [1] 申し込みは、老人課へ申請書を提出していただきます。

文種別：連絡先

- [1] 老人課 (99) 9999

文種別：リンク

- [1] 行政サービス

【図17】

本発明の第1の実施例のサービス重要語 抽出部により抽出された結果の例

・商品と商品に対応する動詞、動詞に対応する対象者について

文種別	文番号	商品	動詞句	対象者	商品確信度
メニュー	1	行政(事)			1
メニュー	2	福祉(事)			1
タイトル	1	日常生活用具(物)	給付 等		3
本文	1	介護(事)	助ける		0
本文	2	日常生活用具(物)	給付・助成		3
本文	3	所得制限(事)	ある	介護者	1
表	3	自動消化器			1
表	4	ガス漏れ警報機			1
表	6	自動火災感知器			1
表	8	移動感知機			1
表	9	購入費	助成		3
表	11	マットレス			1
表	12	体位変換器			1
手続き	1	申請書	提出		2

・対象者について

文種別	文番号	対象者	対象者確信度
メニュー	3	寝たきり高齢者	1
本文	1	寝たきり高齢者、一人暮らし高齢者、痴呆性高齢者	3
本文	3	介護者	1
表	2	一人暮らし高齢者	1
表	5	一人暮らし高齢者	1
表	7	痴呆性高齢者	1
表	10	寝たきり高齢者	1

・手続きについて

文種別	文番号	商品	動詞句	対象者	場所格	手続き確信度
手続き	1	申請書	提出		老人課	3

・連絡先について

文種別	文番号	連絡先部課名	電話番号	住所	連絡先確信度
連絡先	1	老人課	99-9999		3

【図18】

本発明の第1の実施例の重要度設定の例

・商品と商品に対応する動詞、動詞に対応する対象者について

文種別	文番号	商品	動詞句	対象者	商品確信度	重要度
メニュー	1	行政(事)			1	1
メニュー	2	福祉(事)			1	1
タイトル	1	日常生活用具(物)	給付 等		3	9
本文	1	日常生活用具(物)	給付・助成		3	6
本文	2	所得制限(事)	ある		1	1
表	3	自動消化器(物)			1	1
表	4	ガス漏れ警報機(物)			1	1
表	6	自動火災感知器(物)			1	1
表	8	移動感知機(物)			1	1
表	9	購入費(物)	助成	介護者	3	3
表	11	マットレス(物)			1	1
表	12	体位変換器(物)			1	1
手続き	1	申請書(物)	提出		2	2

・対象者について

文種別	文番号	対象者	対象者確信度	重要度
メニュー	3	寝たきり高齢者	1	1
本文	1	寝たきり高齢者、一人暮らし高齢者、痴呆性高齢者	3	6
本文	3	介護者	1	2
表	2	一人暮らし高齢者	1	1
表	5	一人暮らし高齢者	1	1
表	7	痴呆性高齢者	1	1
表	10	寝たきり高齢者	1	1

・手続きについて

文種別	文番号	商品	動詞句	対象者	場所格	手続き確信度	重要度
手続き	1	申請書(物)	提出		老人課	3	3

・連絡先について

文種別	文番号	連絡先部署課名	電話番号	住所	連絡先確信度	重要度
連絡先	1	老人課	99-9999		3	3

【図19】

本発明の第1の実施例の要約語選択制御部による
概念統一、マージ、共通概念の統一処理の例

・商品と商品に対応する動詞、動詞に対応する対象者について

商品	マージした商品	動詞句	対象者	重要度
行政(事)				1
福祉(事)				1
日常生活用具(物)	ガス漏れ警報機、自動火災感知機、移動感知機、自動消化器、マットレス、体位変換器(物)	給付・助成		21
所得制限(事)		ある		1
購入費		助成	介護者	3

・対象者について

対象者	重要度
寝たきり高齢者、一人暮らし高齢者、痴呆性高齢者	11
介護者	2

・手続きについて

文種別	文番号	商品	動詞句	対象者	場所格	重要度
手続き	1	申請書(物)	提出		老人課	3

・連絡先について

文種別	文番号	連絡先部課名	電話番号	住所	重要度
連絡先	1	老人課	99-9999		3

【図20】

本発明の第1の実施例の要約語選択制御部による
重要度の高いレコードの選択処理の例

・商品と商品に対応する動詞、動詞に対応する対象者について

商品	マージした商品	動詞句	対象者	重要度
日常生活用具（物）	ガス漏れ警報機、自動火災感知機、移動感知機、自動消火器、マットレス、体位変換器（物）	給付・助成		21

・対象者について

対象者	重要度
寝たきり高齢者、一人暮らし高齢者、痴呆性高齢者	7

・手続きについて

商品	動詞句	対象者	場所格	重要度
申請書（物）	提出		老人課	3

・連絡先について

連絡先部課名	電話番号	住所	重要度
老人課	99-9999		3

【図21】

本発明の第1の実施例の要約語 選択制御部における情報補完の例

・商品と商品に対応する動詞、動詞に対応する対象者について

商品	マージした商品	動詞句	対象者	重要度
日常生活用具（物）	ガス漏れ警報機、自動火災感知機、移動感知機、自動消火器、マットレス、体位変換器（物）	給付・助成	寝たきり高齢者、一人暮らし高齢者、痴呆性高齢者	21

・対象者について

対象者	重要度
寝たきり高齢者、一人暮らし高齢者、痴呆性高齢者	7

・手続きについて

商品	動詞句	対象者	場所格	重要度
申請書（物）	提出		老人課	3

・連絡先について

連絡先部課名	電話番号	住所	重要度
老人課	99-9999		3

【図23】

本発明の第3の実施例の種別付与・重要語抽出結果の例

(a)	<p>本文1から抽出された内容</p> <p>①商品：介護（事） 動詞：困難 対象者：痴呆性高齢者 重要度：6 ②商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 時間：日中 重要度：6 ③商品：負担 動詞：軽減 対象者：介護者 重要度：6</p>
(b)	<p>本文2から抽出された内容</p> <p>④商品：介護（事） 動詞：困難 対象者：痴呆性高齢者 重要度：6 ⑤商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 時間：夜間 重要度：6 ⑥商品：負担 動詞：軽減 対象者：介護者 重要度：6</p>

【図25】

本発明の第4の実施例の種別付与・重要語抽出された結果の例

(a)	<p>本文1から抽出された内容</p> <p>①商品：介護（事） 動詞：困難 対象者：痴呆性高齢者 重要度：6 ②商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 時間：日中 重要度：6 ③商品：負担 動詞：軽減 対象者：介護者 重要度：6</p>
(b)	<p>本文2から抽出された内容</p> <p>④商品：介護（事） 動詞：困難 対象者：痴呆性高齢者 重要度：6 ⑤商品：痴呆性高齢者（人） 動詞：預かる 場所：老人ホーム 時間：夜間 重要度：6 ⑥商品：負担 動詞：軽減 対象者：介護者 重要度：6</p>